**Section 6.4: Resampling Methods in Python**

**Duration:** 1 hour 45 mins

**Concepts:**

* The validation set approach
* Leave-one-out cross-validation
* K-fold cross-validation
* The bootstrap

**Textbook section:** An Introduction to Statistical Learning on Python, Chapter 5

|  |  |
| --- | --- |
| **Materials and Resources** | **Learning Goals** |
| * Computers for students with Jupyter Notebook * Resampling Methods Slides * Resampling Methods Exercises Jupyter Notebook file | * Learn how to use cross-validation and the bootstrap to find the best model. |

|  |  |  |
| --- | --- | --- |
| **Duration** | **Lesson Section** | **Learning Objectives** |
| 8 mins | Go through the validation set approach of the slides. | * The validation set approach * Drawbacks to this approach |
| 15 mins | Go through “The Validation Set Approach” section in the Jupyter Notebook file as a class. | * Use the validation set approach for a linear model |
| 5 mins | Go through the leave-one-out cross-validation section of the slides. | * Leave-one-out CV * How is it better than the set approach |
| 20 mins | Go through the leave-one-out cross-validation section in the Jupyter Notebook file as a class. | * Use `cross\_validate()` to perform LOOCV for a linear model. * Use `cv\_error` to choose the best degree of polynomial to fit to the data |
| 8 mins | Go through the k-fold cross-validation section of the slides. | * K-fold CV * LOOCV vs k-fold CV |
| 15 mins | Go through the k-fold cross-validation section in the Jupyter Notebook file as a class. | * Use `cv\_error` to choose the best degree of polynomial to fit to the data using k-fold CV |
| 8 mins | Go through the bootstrap section of the slides. | * The bootstrap |
| 15 mins | Go through the bootstrap section in the Jupyter Notebook file as a class. | * Use `boot\_SE()` to find the SE of the mean of a data frame * Use `boot\_OLS()` to find the SE of parameters from lm. |